

Product Unbundling in the Travel Industry: The Economics of Airline Bag Fees

by

Jan K. Brueckner
Department of Economics
University of California, Irvine

Darin N. Lee
Compass Lexecon

Pierre M. Picard
Department of Economics
University of Luxembourg

Ethan Singer
Department of Economics
University of Minnesota

Revised February 2014

Abstract

This paper provides theory and evidence on airline bag fees, offering insights into a real-world case of product unbundling. The theory predicts that an airline's fares should fall when it introduces a bag fee, but that the full trip price (the bag fee plus the new fare) could either rise or fall. The empirical evidence presented in the paper provides strong confirmation of this prediction. The data also suggest that the average fare falls by less than the bag fee itself, so that the full price of a trip rises for passengers who choose to check bags.

Product Unbundling in the Travel Industry: The Economics of Airline Bag Fees

by

Jan K. Brueckner, Darin N. Lee, Pierre M. Picard, and Ethan Singer*

1. Introduction

Product bundling and unbundling by firms has been a focus of researchers in industrial organization ever since the seminal contributions of Stigler (1963) and Adams and Yellen (1976).¹ While consumers may have only been vaguely aware of the impact of firm bundling decisions on the product choices available to them, recent developments in the airline industry have brought the bundling issue into sharp focus. In recent years, airlines have unbundled key elements of the product they offer, introducing separate charges for in-flight food and for checked baggage, services that were previously included in the ticket price. In addition to being obvious to millions of fliers, this product unbundling has received widespread coverage in the press, with the reports typically highlighting passenger unhappiness with the new bag fees.²

The new fees, especially the charges for checked baggage, have generated a substantial amount of revenue for the airlines. Data provided by Bureau of Transportation Statistics (2012) show that airlines collected approximately \$2.5 billion in bag fees in 2009 and \$3.5 billion in 2012. While the typical observer might view this new source of revenue as a windfall for airlines, economists would think twice before reaching such a conclusion. The typical \$20 fee charged for the first checked bag represents an increase in the effective fare for those passengers who pay it, amounting to a 10 percent increase over a typical fare of \$200, measured on a one-way basis. If airlines lack the pricing power to impose such an increase in the cost of travel, economic theory suggests they would have had to cut base fares in offsetting fashion when imposing “first bag” fees.³

The possibility that the adoption of bag fees may put downward pressure on base fares seems to have been entirely overlooked in the voluminous popular discussions of the new fees.

Airlines, of course, never announced any bag-fee-related fare reductions, but the carriers have ample opportunity to quietly reduce fares via their yield management systems by allowing more tickets in the cheapest fare classes to be sold. The purpose of the present paper is to explore this possible connection between airline fares and bag fees, both theoretically and empirically. The paper begins by developing a theoretical model that portrays an airline's profit-maximizing choices of both the base fare and the bag fee. The model predicts a decline in the base fare. Although this fare decline can be viewed as a natural outcome, establishing the result theoretically is not straightforward. The prediction is then tested empirically, using detailed fare data from the US Department of Transportation. The empirical results confirm the theoretical prediction, with an airline's *average* fare falling when it starts collecting a bag fee, but by less than the amount of the fee.⁴ Given the rarity of empirical work on the effects of product unbundling, these empirical results are highly noteworthy. In addition, the findings are distinguished by their focus on a visible unbundling event.

The fact that the checked-bag service would only be purchased conditional on purchase of an airline ticket (never being bought on its own) makes a model of "add-ons" the appropriate analytical framework. Ellison (2005) developed an add-on model, using internet service in a hotel room or extended warranties for cars and appliances as examples, but he assumed that the consumer only learns the price of the add-on when purchasing the basic product. While some airline customers may have been taken by surprise by the presence of bag fees when they were newly introduced, almost everyone now seems to be aware of these charges and their magnitude. The add-on model closest to ours is the simple perfect-information setup proposed by Fruchter, Gerstner and Dobson (2011), which has unitary purchases and two types of consumers.

For simplicity, our model follows their framework by portraying the decisions of a monopoly firm. This airline faces customers with heterogeneous inconvenience costs from *not* checking a bag, which necessitates the use of carry-on luggage. Passengers undertaking long vacations, who require substantial luggage capacity, would have high values of this inconvenience cost, while business passengers making quick trips would have lower values. A bag fee will then deter passengers with low idiosyncratic costs from checking a bag, while others will pay the

fee. Finally, the airline incurs a bag handling cost for each checked bag.

The analysis confirms the familiar conclusion, established earlier by Mussa and Rosen (1978) and Verboven (1999), that the higher-quality product (an airline seat with checked-bag service) is provided at a higher effective price (inclusive of the bag fee) to those passengers who value it most, while the remaining passengers get the lower-quality product (a seat without bag service) at a lower price. From an empirical perspective, however, the model has an important, but natural, new implication: *the fare falls when a bag fee is imposed*. The fare plus the bag fee, however, could be either higher or lower than the pre-bag-fee fare, depending on the shape of the demand curve. Although the model is developed in the context of bag fees, it applies to any unbundling of a product add-on that has heterogeneous valuations across consumers (hotel internet service, extended product warranties). The model therefore represents a new contribution to the theoretical literature on add-ons.

Beyond this contribution, the model serves to motivate the paper's empirical work. This motivation is imperfect, however, because the model's focus on the monopoly case (maintained for reasons of theoretical tractability) means that the results may not be directly relevant for competitive airline markets. However, the prediction that fares fall with imposition of a bag fee is natural and likely to be robust, making it a plausible empirical hypothesis for any market structure.

The paper's empirical work combines DOT fare data, which is available quarterly, with information on the various dates when airlines first adopted their bag fees, also making use of information on subsequent changes in the fee levels. The focus is only on the fee for the first checked bag, which is by far the most prevalent type of fee (fees for second checked bags or overweight bags are not considered). The regression model builds on a recent study by Brueckner, Lee and Singer (BLS, 2013), which estimates standard reduced-form log fare regressions that capture the fare impacts of a wide variety of different types of airline competition. This very detailed fare model is supplemented by information on the adoption of airline first bag fees. A new time-dependent variable is added to the model indicating whether the observed carrier had a first bag fee in place during the given quarter. The fact that bag fees were not adopted simultaneously by the airlines, with adoptions instead spread over a period of 13

months (and absent entirely for two carriers), allows their fare impact to be identified. While this base model does not capture the level of the bag fee, a second model includes different bag fee variables for different fee levels, distinguishing the early \$15 level chosen by most carriers from the \$20 level that became common later. In effect, our regressions can be viewed as hedonic price functions for fares, allowing fare levels to differ depending on whether checked-bag service is included. The sample period covers the range of time over which the first bag fees were adopted, running eight quarters from 2008-Q1 to 2009-Q4.

It is important to recognize that the goal of the empirical work is to measure the effect of bag-fee adoption on fares, not to investigate the forces leading to the imposition of bag fees. Accordingly, the adoption of bag fees is treated as exogenous, being viewed as an institutional innovation that was embraced almost simultaneously by all legacy carriers in response to rising fuel prices and other pressures on profit. Potential concerns about endogeneity bias are lessened by the presence of airline and quarter fixed effects, which control for the intertemporal forces that led to bag-fee adoptions as well as any carrier-specific factors that may have governed the small variations in adoption timing. With these variables included, we expect that unobservable factors correlated with bag-fee adoptions (which could create bias) are largely absent.

The regression results indicate that the average fare charged by an airline declines by about 3 percent when the bag fee is adopted. When this decline is less than the amount of the bag fee, the average fare plus the bag fee rises. Moreover, because the decline in average base fares varies with route characteristics, the decline could conceivably be greater than the amount of the bag fee for some passengers. Moreover, passengers who sometimes, but not always, check bags may have a lower total outlay (including bag fees) over multiple trips if the aggregate saving in base fares is greater than the total bag fees paid. Additional evidence suggests that the fare impact is most prevalent in fares outside the top end of the fare distribution. In providing these findings, the paper offers rare insights into the impact of product unbundling in a major industry.⁵

Section 2 of the paper provides the theoretical analysis, while section 3 describes the empirical framework. Section 4 presents the empirical results, and section 5 offers conclusions.

2. Model

2.1. Basics

To generate a simple model, suppose that the cost of carrying a passenger is a constant c , and that the bag-handling cost is k . Let b denote the bag-fee and p denote the airfare. For simplicity, the existence of different fares targeted at different classes of passengers (business and leisure) is suppressed in the model, with the discussion returning to this issue once results have been derived.

The individual demand functions for passengers is given by $D(t)$, where t is the “full price” of a trip. This function gives the number of trips a passenger is willing to make during a year. The full-price elasticity of demand is given by $\epsilon < 0$, which is in general a function of the full prices. The full price that enters the demand function equals the relevant fare plus any baggage-related cost. For a passenger who checks a bag, this cost is simply the bag fee b . For the remaining passengers, the baggage cost equals the *inconvenience cost of not checking a bag*, denoted $\delta \geq 0$. The full price depends on whether a bag is checked, a decision that depends on the relation between δ and b . If $\delta > b$ holds, then the passenger’s inconvenience cost of not checking a bag exceeds the bag fee, making it optimal to check a bag. Conversely, if $\delta \leq b$, the bag fee exceeds or equals the inconvenience cost of using carry-on luggage, making it optimal not to check a bag. Therefore, the full price of a trip for a passenger with inconvenience cost δ is $t = p + \min\{\delta, b\}$, with demand specified accordingly.

Consider the pre-bag-fee era, where bag fees did not exist. Under the maintained assumptions, all passengers would have checked bags, with passengers having zero values of δ , who are indifferent between checking and not checking, assumed to take the first action.⁶ Letting N denote the total number of passengers, the airline’s profit is then $(p - c - k)ND(p)$, where $p - c - k$ is profit per trip. The pre-bag-fee fare, denoted \tilde{p} maximizes this profit expression, satisfying

$$D(\tilde{p}) \left[1 + \frac{\tilde{p} - c - k}{\tilde{p}} \epsilon(\tilde{p}) \right] = 0, \quad (1)$$

where the price dependence of the elasticity is recognized. For the second-order condition to be satisfied, the expression in (1) must be decreasing in \tilde{p} .

2.2. Main analysis

Suppose now that the airline considers charging a bag fee. The fee may deter some passengers from checking a bag, and this potential outcome may make it optimal to impose a fee. Let $f(\cdot)$ denote the density of δ among passengers, which is assumed to be continuous, and $F(\cdot)$ denote the cumulative distribution function. Without loss of generality, the support of the density is assumed to be $[0, \bar{\delta}]$, so that some passengers have arbitrarily small δ 's. Profit now equals

$$\pi \equiv \int_0^b f(\delta)(p - c)D(p + \delta)d\delta + [N - F(b)](p + b - c - k)D(p + b). \quad (2)$$

Note in (2) that $N - F(b)$ passengers have values of the inconvenience parameter δ above b and therefore check a bag. For those not checking a bag, the full price is $p + \delta$, which appears as the argument of the demand function in the integral. Note that when $b = 0$, (2) reduces to the previous profit expression.⁷

Since a bag fee covers the costs created by bag-checking passengers while deterring some individuals from checking a bag, the airline has an incentive to impose it. Using Leibniz' rule along with $f(b) = F'(b)$, and then canceling terms, differentiation of (1) with respect to b yields

$$-(b - k)f(b)D(p + b) + [N - F(b)]D(p + b) \left[1 + \frac{p + b - c - k}{p + b} \epsilon(p + b) \right] = 0. \quad (3)$$

The first term in (3) is the change in profit from the group of marginal passengers, those who stop checking a bag in response to an incremental increase in the bag fee. Since they no longer pay the fee, revenue for each of these marginal passengers falls by b , while their bag-handling cost disappears, leading to a gain of k per passenger. The resulting profit change per marginal passenger, $-(b - k)$, is multiplied by their number, $f(b)$, times trips per passenger, $D(p + b)$. Note that this profit change is positive when b is small but negative when b exceeds k .

The second term in (3) is the change in the profit earned from inframarginal bag-checking passengers, who make $[N - F(b)]D(p + b)$ trips. Since the higher bag fee reduces individual

demand but raises revenue per passenger, this term can take either sign, depending on the level of the full price.

Differentiating (2) with respect to p , the first-order condition for fare is⁸

$$\int_0^b f(\delta)D(p+\delta) \left(1 + \frac{p-c}{p+\delta}\epsilon(p+\delta)\right) d\delta + [N - F(b)]D(p+b) \left(1 + \frac{p+b-c-k}{p+b}\epsilon(p+b)\right) = 0. \quad (4)$$

The first term is the marginal increase in profit from non-bag-checkers when p increases, while the second term is the marginal profit increase from passengers who check bags. Using (3) and (4), the following conclusion can be derived:

Proposition 1. *If some passengers have zero idiosyncratic costs of not checking a bag, then imposition of a bag fee is optimal.*

This result follows because conditions (3) and (4) rule out a solution with $b = 0$ when $f(0) > 0$, indicating that some passengers incur no inconvenience cost from not checking a bag. If $b = 0$, the second line of (4) must equal zero, implying a zero value for the corresponding expression in (3). But with $f(0) > 0$, the first term in (3), and thus the entire expression, is then positive, implying the desirability of increasing b above zero. The proposition says that if there exists a mass of passengers who will be deterred from checking a bag when the fee increases marginally above zero, then such an increase is optimal.⁹

Subsequent conclusions depend on whether the bag fee is smaller or larger than the bag-handling cost k , and the following result is an ingredient to deciding which case holds:

Lemma. *If the bag fee b exceeds the bag handling cost k , then the full price $p + b$ paid by bag-checkers is less than the pre-bag-fee fare, \tilde{p} . If $b < k$ holds instead, then $p + b > \tilde{p}$.*

The lemma, which is proved in the appendix along with subsequent results, can be understood as follows. When the bag fee rises incrementally, (3) requires that the change in profit from the group of marginal passengers is exactly offset by change in profit from inframarginal

passengers. The profit change for marginal passengers has the sign of $-(b - k)$, while the sign of the profit change for inframarginal passengers depends on whether the full price is larger or smaller than \tilde{p} from (1). For the two profit changes to have opposite signs, the relationships in the lemma must hold. The lemma does not say whether b is larger or smaller than k , and the analysis now turns to this question, showing that the answer depends on the curvature of the demand function.

2.2.1. Constant or increasing elasticities

Consider first the case where the price elasticity is constant or increasing in the full price $t = p + b$, so that $0 > \epsilon(t') \geq \epsilon(t)$ for $t' > t$. With a constant or increasing elasticity, the appendix shows the bag fee must exceed the bag handling cost, so that $b > k$. This conclusion, along with lemma, yields

Proposition 2. *When the demand elasticity is constant or increasing in the full price, the bag fee b exceeds the bag handling cost k . From the lemma, the full price paid by bag-checkers is then less than the pre-bag-fee fare ($p + b < \tilde{p}$), so that the fare falls after adoption of the bag fee while remaining above the cost of transporting a passenger ($c < p < \tilde{p}$).*

The proposition may at first appear surprising because it shows that all passengers pay less for travel after imposition of the bag fee. This outcome, however, is consistent with higher profit because the airline no longer incurs bag handling costs for some passengers.

The proposition also yields a striking welfare result. Since $p + b < \tilde{p}$ holds, all bag checkers are better off than prior to adoption of the bag fee. In addition, since $\delta < b$ holds among non-bag-checkers, it follows that $p + \delta < p + b < \tilde{p}$ holds as well, indicating that the full cost of travel is lower for this group after adoption of the bag fee than before. Therefore, all passengers benefit from adoption of the fee. Since airline profit rises as well, the following conclusion emerges:¹⁰

Corollary. *When the demand elasticity is constant or increasing, social welfare rises with adoption of the bag fee. The increase in airline profit is accompanied by a welfare gain for all passengers.*

2.2.2. Linear demand

A general analysis of the case where the demand elasticity is decreasing in the full price appears to be infeasible. But results for the linear case, which has this property, are available:¹¹

Proposition 3. *When the demand function is linear, the bag fee b is less than the bag handling cost k . Although, from the lemma, the full price paid by bag-checkers is then greater than the pre-bag-fee fare ($p + b > \tilde{p}$), the fare falls after adoption of the bag fee while remaining above the cost of transporting a passenger ($c < p < \tilde{p}$).*

In addition to Proposition 3, a further argument shows that $p < \tilde{p}$ holds for any demand function that makes $b < k$ optimal.¹² Together with the lemma, this result establishes that $p < \tilde{p}$ is always satisfied.¹³

Since $p + b > \tilde{p}$ holds in the linear case, bag-checkers are worse off than before adoption of the bag fee. In addition, the welfare impact on non-bag-checkers is ambiguous. It is negative for those with high inconvenience cost (δ close to b) and positive for the others. Thus, in contrast to the earlier analysis, the impact of introduction of the bag fee on social welfare is ambiguous.

2.4. Numerical examples

To illustrate the propositions and get a sense for numerical magnitudes, it is useful to generate some numerical examples using both a constant-elasticity demand function, which takes the form $D(t) = t^\epsilon$, and a linear demand function, $D(t) = \tau + \mu t$. In the examples, the cost c of transporting a passenger is set at \$100 (a number that seems realistic for a typical one-way trip), and the bag handling cost k takes the values of \$5, \$10 and \$20. The distribution of δ is assumed to be uniform over the interval $[0, 30]$. The second-order conditions hold at all the solution values.

The upper panel of Table 1 shows the results for constant elasticity demand with the elasticity ϵ set at values of -3 and -2 . Several regularities are apparent. First, while the bag fee is larger than k , conforming to Proposition 2, the differences are small, all being less than 50 cents. In addition, while the full price $p + b$ for bag checkers lies below the pre-bag-fee fare \tilde{p} , as predicted, the difference is usually small, although it approaches \$7 when $k = 20$ and $\epsilon = -2$. Thus, the fare falls by somewhat more than the fee itself. While less-elastic demand

naturally raises the fare, there is little effect on gap between b and k , which falls by only about 5 cents. These patterns are robust to changes in c and to the use of other simple δ density functions (triangular, linear increasing, linear decreasing).

The lower panel of Table 1 shows the results for linear demand, with the intercept τ set at 210 and the slope μ taking values of -1.0 and -0.6 . While the bag fee is again almost the same as k , it is now smaller rather than larger, as predicted by Proposition 2. In addition, the full price $p + b$ is now larger than the pre-bag-fee fare, as predicted. As in the constant-elasticity case, the gap is usually small, although it approaches \$3 when $k = 20$ and $\mu = -0.6$.

Since the numerical results show that b is close to k , generating bag fees in the observed \$15–20 range requires a large value for the bag handling cost. While values of the required magnitude at first might seem implausible, several additional considerations could suggest otherwise, as follows.

The labor cost of handling a single bag on a one-way trip probably is no larger than than a few dollars, but this number fails to include the capital cost of the airline’s baggage-handling infrastructure, which would inflate the magnitude. In addition, since checked bags are larger and heavier than carry-on bags, they require more fuel to transport, providing another increment to k . Finally, k may include the opportunity cost of aircraft cargo space that is occupied by a checked bag rather than lucrative air freight. Since air cargo earns about \$1 per kilogram, accommodating a 20-kilogram checked bag requires sacrificing \$20 in revenue, implying a k value capable of generating realistic bag fees under the model.¹⁴

2.5. Empirical implications

The preceding results show that, regardless of the form of the demand function, adoption of the bag fee leads to a reduction in the fare. This is the main prediction of the analysis, yielding

Empirical hypothesis: *Other things equal, fares are lower after the adoption of bag fees than before.*

As noted in the introduction, this hypothesis, having been derived in a monopoly context, may not be directly relevant as a prediction applied to competitive airline markets. However, a

competitive version of the model would be difficult to develop, and a monopoly-based prediction provides at least some guidance for the empirical work, especially since the prediction is a natural one that presumably would hold under other market structures. In addition, the model generating this prediction has independent value as a contribution to the theoretical literature on add-ons.

Another issue is the model's suppression of different classes of passengers. In reality, airlines attempt to segment the market for air travel into leisure and business submarkets, in which different fares can be charged. The model could be adapted to handle such price discrimination by introducing two groups with different price elasticities and different inconvenience-cost distributions. The airline would then solve separate profit-maximization problems for each group. However, this approach would generate group-specific bag fees, which could be viewed as unrealistic. On the other hand, since purchase of a full-fare ticket usually exempts the passenger from the bag fee, and since elite frequent-fliers are often exempt as well, the airlines in effect levy a zero bag fee on some business passengers, although others (who purchase cheaper tickets or lack elite status) are still subject to fees. To the extent that business passengers face bag fees less often than leisure passengers and are reimbursed for such fees, the analysis suggests that business fares may be less affected than leisure fares by the imposition of fees.

3. Empirical Framework

As explained in the introduction, the recent paper by Brueckner, Lee and Singer (2013) offers a useful framework for empirically estimating the effect of bag fees on airfares. BLS builds on many of the airline pricing studies from the 1980s and 1990s by incorporating recent industry developments, particularly competition from low cost carriers (LCCs) such as Southwest, JetBlue and AirTran.¹⁵ Given this focus, the BLS framework is particularly well-suited to exploring the impact of bag fees on fares, since any potential price impact would need to be disentangled from both the competitive environment in a particular airline market (the presence of LCCs, number of legacy carriers¹⁶) as well as other market characteristics (route distance, demographic factors).

To adapt the basic BLS model to present purposes, we introduce two primary changes.

First, the BLS sample period, which consists of the four quarters ending 2008-Q2, is expanded to cover the period 2008-Q1 through 2009-Q4, thereby capturing the period over which bag fees were introduced. Second, a bag fee variable is added to the model, indicating whether a fee for a first checked bag was in place for the given carrier during the quarter. This variable equals zero for a given carrier in the quarters prior to the introduction of its bag fee and one for quarters following the quarter of introduction. For the quarter when the bag fee is introduced, the variable equals the fraction of the quarter during which it was present, a structure that accommodates the quarterly aggregation of the fare data (see below). To separate the effect of bag fee introductions from time trends in fares, the model also includes quarter dummies.

First bag fees were implemented during the second half of 2008 and the first part of 2009, and their introduction dates by carrier are as follows: United (June 13, 2008), American (June 15, 2008), US Airways (July 9, 2008), Northwest (August 28, 2008), Sun Country (October 1, 2008), Continental (October 7, 2008), Midwest (October 21, 2008), Frontier (November 1, 2008), Delta (December 5, 2008), AirTran (December 5, 2008), Virgin America (May 5, 2009), and Alaska (July 7, 2009). Spirit and Allegiant both began charging fees for a first bag prior to the start of the sample period. Southwest and JetBlue do not currently charge for bags.¹⁷

It should be noted that charges for checked bags beyond the first bag were in place at most carriers before the start of the sample period. Therefore, a control for the presence of these second-and-additional bag fees is not needed in the model.

The data source for the analysis is the USDOT's DB1B database, a 10 percent quarterly sample of all domestic airline tickets that serves as the primary data source for almost all empirical studies of pricing in the US airline industry. In the raw DB1B data, an observation contains the itinerary's routing (the origin and destination airport, as well as any intermediate connecting points), the operating and marketing carrier for each flight segment, and the fare, including taxes and airport (passenger facility) charges. The DB1B data does not include any airline fees such as bag fees that are not part of the base fare.

We apply various standard filters to the raw data in generating our data set, and all fares are expressed on a one-way equivalent basis, with the inbound and outbound parts of round trips counted as separate observations (ignoring their directionality).¹⁸ We also follow BLS

in equating airline markets with airport-pairs instead of using a city-pair approach, under which airports in a multiple-airport metro area are aggregated into a single destination. After applying our filters and using the airport-pair market definition, we aggregate the data up to the market-carrier-quarter level. The dependent variable is then the log of the passenger-weighted average fare for a particular carrier in a given market during a particular quarter.¹⁹

Finally, the sample is split into two different types of markets: “nonstop” markets, where nonstop service is provided by at least one legacy carrier, and “connect” markets, which do not enjoy any nonstop service from legacies or LCCs, with passengers needing to make a connection to travel in the market. The rationale for this division is that pricing patterns may differ across the two types of markets. Note that nonstop markets may also have connecting service.²⁰

The BLS model is distinguished by its measurement of the fare impacts from a wide variety of different types of competition, and we retain this approach.²¹ In nonstop markets, legacy nonstop competition is measured by the two variables *leg_ns2* and *leg_ns3*, whose coefficients give the fare impacts of adding second and third legacy nonstop carriers to the market (at least one is already present by construction).²² LCC nonstop competition is measured by two dummy variables, *D(WN_ns)* and *D(otherlcc_ns)*, which indicate nonstop competition by Southwest (whose code is WN) or one of the other LCCs.²³ LCC competition in many markets often comes from service at adjacent airports (Midway in Chicago, Baltimore-Washington in Washington, D.C.), and the dummy variables *D(WN_adj_ns)* and *D(otherlcc_adj_ns)* are designed to capture such service. They equal one for a given airport-pair when Southwest or another LCC provides service between the airport-pair’s two cities using adjacent airports at one or both endpoints. Adjacent competition by legacy carriers is discussed in the appendix. See BLS (2013) for the list of airports considered to be adjacent to one another in multiple-airport metro areas.

In the nonstop markets, competition from connecting service is also counted. The legacy competition measure (which counts the number of connecting competitors) is discussed in the appendix, while LCC connecting competition is captured by the dummy variables *D(WN_connect)* and *D(otherlcc_connect)*. Potential competition from LCC carriers, which exists when LCCs serve the market endpoints without actually providing service in the market itself,

is also captured through additional variables, as explained in the appendix.

The treatment of connect markets differs from that of nonstop markets mainly through the absence of all the variables that measure nonstop competition. Legacy competition, which is necessarily of the connecting type, is captured by the variables *leg_connect2*, *leg_connect3* and *leg_connect4m*, which respectively give the fare impacts of adding a second, third, and additional competitors beyond three to the market.

The regressions include carrier fixed effects to control for cost differences across airlines. While the regressions also include quarter dummies, the presence of only eight quarters of data prevents effective use of market fixed effects to capture the influence of time-invariant market characteristics on fares. Instead, several characteristics measures are created, following the literature. The first is *ltdist*, a log distance measure for the market.²⁴ The second variable is *pop*, equal to the geometric mean of the endpoint city populations, while the third is *income*, equal to the geometric mean of the per capita incomes of the endpoint cities. Population and income affect the demand for travel in airline markets, with higher demands possibly raising fares. The demand effect of higher income may arise from a greater propensity for personal air travel as well as from a greater share of high-demand, time-sensitive business passengers in a market. However, the positive demand effects on fares from both variables could be offset (at least partially) through cost savings associated with higher traffic densities, which are realized when high demand leads to large passenger volumes on a route.²⁵

A fourth market-characteristics variable is *tempdiff*, the absolute difference between the average January high temperatures at the endpoint cities. A high value of this variable is likely to indicate a leisure market, where vacation passengers travel from cold to warmer climates. With leisure travel demand being more elastic, a high value of *tempdiff* would be expected to generate lower fares. An additional itinerary-level variable is the dummy $D(\text{connect})$, which indicates that the observation pertains to connecting service in a nonstop market.²⁶ Since connecting trips are less convenient than nonstop trips, the expected fare impact of $D(\text{connect})$ is negative.

The last and most important variable for present purposes is *bagfee*, which captures the presence of a bag fee during the quarter pertaining to the fare observation. Summary statistics

for the nonstop and connect samples are shown in Table 2, which gives means weighted by the passenger count for each market-carrier-quarter cell. Note that the mean of *bagfee* is 0.5, indicating that the half of the sample passengers flew on a carrier that charged a bag fee at the time of their travel. Note that, like Table 2, the regressions reported below also reflect passenger weighting. The sizes of the nonstop and connect samples are 65,554 and 117,705, respectively.²⁷

As explained in the introduction, the adoption of bag fees is treated as exogenous, being viewed as an institutional innovation that was embraced almost simultaneously by all legacy carriers in response to rising fuel prices and other factors affecting profit. For this treatment to be appropriate, important sources of correlation between the *bagfee* variable and the regression error term must be absent. Industry-wide forces such as escalating fuel prices that drove bag-fee adoptions are captured by the time dummies in the regressions, eliminating one possible source of such correlation. While there is little hope of fully explaining the particular order in which carriers adopted the fees over the short, eight-quarter time interval, the carrier dummies in the regressions capture carrier-specific factors that may have influenced the particular timing of adoption, thereby eliminating another source of possible correlation. With these controls present, it is hard to identify any remaining factors that are unobservable, affect fares, and are correlated with the presence of bag fees. As a result, treating the *bagfee* variable as exogenous is warranted.

In gauging the likely impact of a bag fee on fares, it is important to recognize that the dependent variable for most of the regressions is the average fare charged by the given carrier in the market. This average collapses into a single value a variety of different fares, including the relatively high fares paid by business travelers unable to book trips far in advance along with the lower fares paid by leisure travelers. As a result, our fare measure will mask what is likely to be substantial variation across fare types in the bag fee's fare impacts. But without more information, the estimated impact on the average fare will not reveal the separate impacts in different fare categories. The key required information, namely, the booking class of the individual observed fares, is not reliably available in the DOT data.²⁸ In addition, as discussed earlier, the average fare impacts could mask substantial differences at the passenger level

between those who on balance pay less for air travel and those who pay more because of the adoption of first bag fees.

While evidence of a decline in the average fare would still provide strong support for the main empirical hypothesis, the analysis also explores three different ways of generating sharper conclusions. The first is to allow the bag fee's fare impact to depend on market characteristics, which is made possible by including interaction terms between *bagfee* and two of the characteristics variables (*tempdiff* and *income*) in the regression. Since the share of leisure passengers should be high in markets with a high value for *tempdiff*, the average-fare impact of a bag fee should be larger in such markets, implying a negative coefficient for the *bagfee_tempdiff* interaction variable. Conversely, since markets with a high value of *income* are likely to contain larger shares of both business passengers and leisure passengers with incomes high enough to purchase bag-fee-exempt tickets, the fare impact of *bagfee* should be smaller in such markets. Thus, a positive coefficient on the *bagfee_income* interaction variable should emerge. In order to facilitate interpretation of these interactions, both the *tempdiff* and *income* variables are measured as deviations from their means.

A second approach to eliciting more information on bag-fee impacts is to recognize that the magnitude of bag fees changed during the sample period for most carriers once the fees were imposed, rising from \$15 to \$20. Accordingly, we run regressions that allow the bag fee's fare impact to differ by the magnitude of the fee. A third approach to generating additional information is to make use of fare percentiles. Prior to aggregating the data, the percentiles of a carrier's fare distribution in each market can be computed. For example, the 25th and 85th percentile fares can be generated for each market-carrier-quarter combination. Although such percentiles do not exactly match up with fare booking classes, the 25th percentile fare is very likely to be a leisure fare while the 85th percentile almost certainly represents a business fare.

4. Empirical Results

4.1. Results without the bag-fee variable

The basic regression results are presented in Table 3. Consider first columns 1 and 2, which show estimates of the model without the bag-fee variable. The results for the nonstop

model, which are similar to those in BLS, show a modest fare impact of legacy competition. Addition of a second nonstop legacy carrier to a market leads to only a 3.5 percent reduction in the average fare, while addition of a third carrier has a counterintuitive positive effect. By contrast, the nonstop presence of Southwest in the market reduces fares by a dramatic 31 percent, while the presence of other LCC carriers leads to a 16 percent fare reduction.²⁹ This sharp contrast between the effects of legacy and LCC competition was the main message of the BLS paper.

While the presence of connecting LCC service has no fare impact, adjacent nonstop competition generates appreciable fare reductions: 17 percent for Southwest and 7 percent for other LCCs. The coefficients for the remaining competition variables are presented and discussed in the appendix. Turning to the other covariates, a longer route distance raises fares, while fares are higher in high-income markets and lower in leisure markets, where *tempdiff* is large. The coefficient of $D(connect)$ is negative, as expected, but it is not significantly different from zero. The quarter dummies show that fares are higher than in base quarter (2008Q1) through the rest of 2008 and lower thereafter.

In the connect model, the addition of a second legacy connecting competitor counterintuitively raises fares, while a third competitor leads to a 2.7 percent fare reduction, with further competitors reducing fares by less than 1 percent. Connecting competition from Southwest leads to a 6 percent reduction in fares, while connecting competition from other LCCs leads to a larger 10 percent reduction, perhaps testifying to the greater hub-and-spoke orientation of these carriers relative to Southwest (which may allow more-convenient connecting service). Adjacent nonstop LCC competition has no fare effect.³⁰

4.2. Basic results

Because the focal competition variables in the BLS model are measured at the market level, the regressions in that study cluster the standard errors at this level, an approach that is also followed in columns 1 and 2 of Table 3. Since the bag-fee variable is measured at the carrier-quarter level, Table 3 compares two versions of the regressions that contain the bag-fee variable, which are based on two different levels of clustering of the standard errors: one using the previous market-level clustering (columns 3 and 4) and one with clustering at the

carrier-quarter level (columns 5 and 6).³¹

The first observation from Table 3 is that since the clustering scheme does not affect the point estimates, the corresponding columns (3 vs. 5, 4 vs. 6) differ only in the coefficient standard errors.³² Second, the *bagfee* coefficients themselves indicate a 2.7 percent average-fare reduction in the nonstop model when a bag fee is present, with the average fare falling by 2.4 percent in the connect model. While both *bagfee* coefficients are significant at the 1 percent level under market-level clustering, the standard errors increase when clustering is performed at the carrier-quarter level, an expected reduction in precision given the fact that the bag-fee variable varies only by quarter and carrier. This change reduces the significance of the bagfee coefficient in the nonstop model to 10 percent (the connect coefficient becomes insignificant). While *bagfee*'s low variability produces this outcome, other specifications explored below make use of the richness in the underlying market-level data by introducing interactions between *bagfee* and market characteristics, by distinguishing between \$15 and \$20 bag fees, and by allowing fare impacts to differ across fare percentiles. As will be seen, these specifications allow finer measurement of the fare impacts of the bag fee.

The percentage impacts from Table 3, which provide some initial confirmation of the maintained hypothesis, can be translated into dollar fare impacts. The fare charged by legacy carriers is the relevant one for this calculation since most LCCs do not charge bag fees. With the legacy average fare in the nonstop model (weighted by passengers) equal to \$196 on a one-way basis, the 2.7 percent impact translates into a nonstop fare reduction of \$5.30. With the average legacy connecting fare equal to \$229 on a one-way basis, the 2.4 percent impact leads to a reduction of \$5.50. This reduction in the average fare could be associated with more-substantial reductions at particular points in the fare distribution. We provide additional evidence on this issue below.

The fact that the point estimates of the dollar drop in the average fare are roughly equal in the nonstop and connect markets may provide information about the sources of the inconvenience of not checking a bag. Although the model assumed that this inconvenience comes from the resulting restriction on a passenger's luggage volume, the distinction between non-stop and connecting travel introduces another possible source of inconvenience: the need to

manage carry-on luggage on a more-complex connecting trip. This added cost would shift the density $f(\cdot)$ of inconvenience costs to the right, with consequent effects on the optimal p and b . However, the fact that the dollar average-fare discount is virtually the same in the nonstop and connect markets suggests this source of inconvenience may not be important.

4.3. Interaction specification

Since we expect fare impacts to differ according to market characteristics (some of which reflect the passenger mix), and since these characteristics exhibit substantial variation within the sample, greater precision can be achieved by allowing the bag-fee effect to depend on market characteristics. Accordingly, the basic regression is supplemented with variables that interact the *bagfee* dummy and the various market-characteristics measures (for example, a *bagfee_tempdiff* interaction variable). For ease of interpretation, the markets-characteristics variables are expressed as deviations from their passenger-weighted means.

The presence of the interaction variables also allows a change in the clustering approach. Since the bag-fee fare impact is now both carrier and market-specific, with the latter dimension introducing most of the heterogeneity given the large number of markets with varying characteristics, the results reported in Table 4 reflect clustering by market, following the original BLS approach. Clustering by carrier-market is another possibility, but it leads to similar qualitative conclusions. The number of markets varies slightly across quarters, with average number of nonstop markets equal to 1,352 and the average number of connect markets equal to 3,659.

The regressions in Table 4 include the uninteracted *bagfee* variable along with the interaction terms. With the interactions constructed using differences from market-characteristic means, the coefficients of the uninteracted *bagfee* variable give the bag fee's fare effect in an average market, where the characteristics take mean values and the interaction variables are thus zero. In contrast to columns 5 and 6 of Table 3, all of these *bagfee* coefficients are negative and significantly different from zero at the 1 percent level, for both nonstop and connect regressions, and regardless of which interactions are included, matching the pattern those in columns 3 and 4 of Table 3.

Turning to the interaction coefficients, columns 1 and 5 present nonstop and connect regressions containing only the *bagfee_tempdiff* interaction. The nonstop interaction coefficient

is negative and significant, as expected, indicating that the bag fee leads to a larger average-fare decrease in leisure markets. This *tempdiff* effect is absent, however, in connect markets, where the coefficient is insignificant. To gauge the size of the *tempdiff* effect, note first that the nonstop coefficient of the uninteracted *bagfee* variable is 0.0265, identical to the coefficient in Table 3. This estimate indicates a 2.7 percent fare reduction from bag-fee imposition in a nonstop market with an average *tempdiff* value. However, in a nonstop market where *tempdiff* exceeds its average by 20, or about two standard deviations, the percentage fare decline from the bag fee is $100(0.0265 + 20 \times 0.0012) = 5.1$ percent, an effect almost double the impact in an average market. Note that the uninteracted *bagfee* coefficient for connect markets in column 6 is also almost identical to that in Table 3, showing a 2.3 percent fare reduction from bag-fee imposition in an average market, but that it is now significant at the 1 percent level.

Columns 2 and 6 show the regressions with the *bagfee_income* interaction variable, whose coefficient is positive and significant for both nonstop and connect markets. Thus, as expected, adoption of a bag fee reduces the average-fare by less in a higher-income market, where more travelers are business passengers or leisure passengers buying full-fare tickets that are exempt from bag fees. According to the nonstop estimate, a \$10,000 increase in income above the mean (which raises *income* by 10) is almost sufficient to reduce the bag-fee effect to zero. Note also that the bag-fee impact in an average-income nonstop market is higher, at 5.1 percent, than in column 2, an outcome that is presumably a consequence of the effect of an asymmetric income distribution combined with the interaction specification. However, the average-market effect in the connect regression, at 2.6 percent, is similar to that in column 5.

For completeness, columns 3 and 7 of Table 4 show a distance-interaction specification. The logic is that the higher fares for longer-distance trips should fall by a smaller percentage when a (distance invariant) bag-fee is adopted. The *bagfee_dist* coefficients, which are positive and significant in both the nonstop and connect models, confirm this expectation. In addition, the bag-fee fare impacts in average nonstop and connect markets are again similar to those in columns 1 and 5. Finally, columns 4 and 8 show regressions that include all the interactions together, and results mostly match those in the specifications where the variables appear singly.

4.4. The effects of different bag-fee magnitudes

While new bag fees were typically instituted at a \$15 level, many carriers raised their fees to \$20 later in the sample period.³³ The different fee levels are not recognized in the regressions discussed so far, but Table 5 makes a distinction between them. In the table, *bagfee15* is a dummy variable indicating the presence of a \$15 fee, while *bagfee20* indicates a \$20 fee. Although carrier-quarter clustering is reintroduced, the greater data heterogeneity introduced by the two bag-fee variables potentially allows satisfactory precision in the measurement of fare effects. In the nonstop model, the coefficient of *bagfee20* is significant at the 5 percent level. In addition, it is more than 50 percent larger in absolute value than *bagfee15*'s coefficient, a difference that is not, however, statistically significant (also, the latter coefficient is only significant at the 10 percent level). The point estimates, however, imply an average-fare drop of \$4.86 in nonstop markets when a \$15 bag fee is adopted, with the discount growing to \$7.86 with a \$20 fee. While this relationship is the expected one given the different magnitudes of the fees, the coefficients have the opposite relationship in the connect model, although neither one is statistically significant.

4.5. Percentile fare regressions

While the interaction approach in Table 4 provides compelling indirect evidence that the fare impact of the bag fee differs across ticket classes, additional evidence comes from yet a different approach that exploits information about the fare distribution underlying a market's average fare, as explained above. Prior to aggregation of individual ticket data, the fare levels corresponding to the 25th, 50th, 75th and 85th percentiles in the fare distribution are tabulated for each market-carrier-quarter combination (Table 2 shows the average fare values at these percentiles). Instead of using a single fare observation, each market-carrier-quarter cell in the first regression in Table 6 is then assigned two fare values, the 25th-percentile fare and the 85th-percentile fare, with the number of observations thus roughly doubling in size.³⁴ The 85th-percentile fare is used rather than a somewhat lower percentile to ensure that few leisure passengers pay it. The regression using this "stacked" data set then includes different intercepts for the two fare percentiles, as well as separate bag-fee variables (*bagfee_p25* and *bagfee_p85*) generated by interacting the fare-percentile dummies with the *bagfee* variable. This regression

is equivalent to running two separate regressions, one for each of the two fare percentiles, with the constraint that all coefficients other than the intercepts and the bag-fee coefficients are equal across the regressions.³⁵

Because the thinness of typical connect markets makes their fare distributions relatively sparse compared to those of nonstop markets, we focus on the nonstop case, using carrier-quarter clustering. Table 6 shows estimates for the nonstop model, and the results in column 1 are striking. They show that the 25th-percentile fare drops by 5.5 percent when the bag fee is adopted, while *bagfee_p85*'s coefficient is insignificant. An equality test shows that the two bag-fee coefficients are significantly different from one another. Note that the *p85* intercept estimate indicates that the 85th-percentile fare is 71 percent larger than the 25-percentile fare.

Column 2 shows the effect of adding the 50th percentile fare to the regression, which results in a near tripling of the data set relative to its original size. The coefficient of *bagfee_p85* remains insignificant, while the 25th percentile fare effect shrinks somewhat to 4.4 percent. In addition, *bagfee_p50*'s coefficient is negative and significant, indicating an intermediate-size 3.9 percent decline in the 50th percentile fare when the bag fee is adopted. While the 25th and 50th-percentile coefficients are not significantly different from one another, both are significantly different from the 85th-percentile coefficient.

The regression in column 3 shows the effect of replacing the 50th-percentile fare with the 75th percentile value. The 25th-percentile effect increases bit to 5.1 percent, while the 85th-percentile coefficient remains insignificant. The 75th percentile effect is of intermediate size at 3.5 percent. Matching the previous case, the 25th and 75th-percentile coefficients are not significantly different from one another, but both are significantly different from the 85th-percentile coefficient.

While the results show that the exact coefficient estimates are somewhat sensitive to the collection of percentiles represented in the regression, the estimates as a whole suggest that the fare impacts of imposing a bag fee indeed differ across fare percentiles, being largest, as expected, at the lower percentiles and smallest at the higher percentiles.

These estimates can be used to give a sharper picture of the magnitude of dollar fare impacts. Applying the 5.8 percent figure from column 1 to the average 25th-percentile legacy

fare of \$128 yields a decline of a bit more than \$7.00 when a bag fee is adopted. Therefore, it appears that leisure fares fall by roughly one-half to one-third of the amount of the bag fee, implying an average increase in the full trip for passengers paying a bag fee that equals about one-half to two-thirds of the fee. As discussed earlier, the impact on individual passengers will depend on the specific fares paid, the route(s) traveled, the number of trips on which the bag fee was paid, whether the passengers was reimbursed for the fee, and other factors. While this qualitative outcome matches the linear-demand case in Table 1, the full-price increase in those examples is not as large, with the average full trip price for passengers paying a bag fee rising by at most one-sixth of the fee.

5. Conclusion

This paper has provided theory and evidence on airline bag fees, offering rare insights into a real-world case of product unbundling. The theory predicts that an airline's average fare should fall when it introduces a bag fee, but that the full trip price for passengers paying the bag fee (the fee plus the new fare) could either rise or fall. The empirical evidence presented in the paper provides strong confirmation of this prediction.

Although measuring the bag fee's effect on the different fare types underlying the average is problematic, the percentile regressions show that the 25th percentile fare falls by about \$7.00 when a bag fee is adopted, an amount equal to about one-half to one-third of the fee. As a result, it appears that the full trip price rises for the average leisure passenger by at least half of the bag fee on trips when that passenger checks a bag. Non-bag-checkers among leisure passengers, however, benefit from a lower fare.

One question raised only briefly so far concerns the timing of this product unbundling by the airlines. Why did the carriers institute bag fees only recently rather than years earlier? Apparently, the explanation lies in the extraordinary recent downward pressure on airline profits, a consequence of 2008's fuel price spike combined with the subsequent severe recession. In trying to at least break even in this harsh environment, airlines instituted what they knew would be an unpopular unbundling of checked-bag services, and the current rosier profit picture may be due in part to this decision.

Appendix

A1. Proof of the lemma

Note first that (3) is satisfied when $p + b = \tilde{p}$ and $b = k$. Then observe that, because the expression in (1) is a decreasing function of p , the second expression in (3) is similarly decreasing in $p + b$. Therefore, when $b > k$, the positive value of the second expression required by (3) in turn requires $p + b < \tilde{p}$. Conversely, $b < k$ requires $p + b > \tilde{p}$.

A2. Proof of Proposition 2

Suppose to the contrary that $b \leq k$. Given (3), the expression

$$1 + \frac{p + b - c - k}{p + b} \epsilon(p + b) \tag{a1}$$

must then be negative (requiring $p > c$ when $b \leq k$, given $\epsilon < 0$). When $\delta \leq b \leq k$, the term inside the integral in (4) satisfies the following inequalities:

$$1 + \frac{p - c}{p + \delta} \epsilon(p + \delta) < 1 + \frac{p + b - c - k}{p + b} \epsilon(p + \delta) \leq 1 + \frac{p + b - c - k}{p + b} \epsilon(p + b), \tag{a2}$$

where the second inequality stems from constant or increasing elasticities. Because the last term is negative, the first term is also negative, which means that (4) cannot be satisfied. This contradiction implies that $b > k$ must hold.

A3. Additional competition variables

Table A1 presents coefficients for additional competition variables from the regressions in Table 3. As mentioned in the text, *nleg_connect* equals the number of legacy carriers providing connecting service in the market (carriers already providing nonstop service are not included in this count). As can be seen, the coefficient of this variable is insignificant. The number of legacy carriers providing nonstop service from adjacent airports is given by *nleg_adj_ns*, and the number providing adjacent connecting service is given by *nleg_adj_connect* (double counting is again avoided). The coefficients of the first variable are insignificant in the nonstop model in Table A1 (though counterintuitively positive and significant in the connecting model), while

the coefficients of the second variable are always significantly positive, again a counterintuitive outcome. The analogous dummy variables $D(WN_adj_connect)$ and $D(otherlcc_adj_connect)$ are computed for LCCs, but they are used only in the connecting model since adjacent connecting service by LCC's is almost nonexistent in the nonstop markets. The coefficients of both these variables are significant in the connecting model, showing a negative fare impact from adjacent competition by LCCs.

Potential competition by LCC carriers is captured by the remaining four variables in Table A1. To be considered a potential competitor, an LCC has to serve 5 or more nonstop routes out of each endpoint airport of a given market, with no LCCs serving the market itself. The two dummy variables indicating such presence are $D(WN_pot)$ and $D(otherlcc_pot)$. Potential LCC competition can also come from adjacent airports, and the dummy variables indicating such a presence are $D(WN_adj_pot)$ and $D(otherlcc_adj_pot)$. As can be seen from the table, in-market potential competition leads to lower fares in the nonstop markets when it comes from Southwest but has no effect when it comes from other LCCs (both types of potential competition reduce fares, though, in the connect markets). Adjacent potential competition matters only in the nonstop case, with all LCCs having an effect.

Table 1

Numerical Examples*Constant-elasticity demand*

| ϵ | k | b | p | $p + b$ | \tilde{p} |
|------------|-----|-------|--------|---------|-------------|
| -3 | 5 | 5.07 | 152.22 | 157.29 | 157.50 |
| -3 | 10 | 10.21 | 153.92 | 164.13 | 165.00 |
| -3 | 20 | 20.39 | 156.00 | 176.39 | 180.00 |
| -2 | 5 | 5.05 | 204.52 | 209.57 | 210.00 |
| -2 | 10 | 10.16 | 208.11 | 218.67 | 220.00 |
| -2 | 20 | 20.30 | 212.63 | 232.93 | 240.00 |

Linear demand

| μ | k | b | p | $p + b$ | \tilde{p} |
|-------|-----|-------|--------|---------|-------------|
| -1.0 | 5 | 4.83 | 152.85 | 157.68 | 157.50 |
| -1.0 | 10 | 9.45 | 151.21 | 160.66 | 160.00 |
| -1.0 | 20 | 18.67 | 148.82 | 167.49 | 165.00 |
| -0.6 | 5 | 4.92 | 222.78 | 227.70 | 227.50 |
| -0.6 | 10 | 9.74 | 221.00 | 230.75 | 230.00 |
| -0.6 | 20 | 19.44 | 218.53 | 237.97 | 235.00 |

Table 2: Summary Statistics

| | Non-Stop Model | | Connect Model | |
|-------------------------|----------------|--------|---------------|--------|
| | mean* | sd | mean* | sd |
| bagfee | 0.501 | 0.485 | 0.576 | 0.477 |
| nleg_ns | 1.427 | 0.617 | | |
| D(WN_ns) | 0.124 | 0.330 | | |
| D(otherlcc_ns) | 0.244 | 0.429 | | |
| nleg_connect | 0.120 | 0.397 | 2.227 | 0.948 |
| D(WN_connect) | 0.045 | 0.207 | 0.202 | 0.401 |
| D(otherlcc_connect) | 0.019 | 0.136 | 0.134 | 0.341 |
| nleg_adj_ns | 0.420 | 0.728 | 0.106 | 0.427 |
| D(WN_adj_ns) | 0.141 | 0.348 | 0.012 | 0.110 |
| D(otherlcc_adj_ns) | 0.140 | 0.347 | 0.017 | 0.130 |
| nleg_adj_connect | | | 0.048 | 0.251 |
| D(WN_adj_connect) | | | 0.004 | 0.062 |
| D(otherlcc_adj_connect) | | | 0.002 | 0.039 |
| D(WN_pot) | 0.089 | 0.284 | 0.211 | 0.408 |
| D(otherlcc_pot) | 0.025 | 0.157 | 0.008 | 0.090 |
| D(WN_adj_pot) | 0.077 | 0.267 | 0.026 | 0.158 |
| D(otherlcc_adj_pot) | 0.023 | 0.149 | 0.007 | 0.081 |
| tdist | 1012.6 | 635.1 | 1452.3 | 665.7 |
| D(connect) | 0.141 | 0.348 | | |
| pop | 4.614 | 2.919 | 1.766 | 1.625 |
| income | 43.541 | 4.905 | 39.533 | 5.187 |
| tempdiff | 15.697 | 10.473 | 16.930 | 10.670 |
| fare | 177.40 | 64.01 | 218.72 | 67.45 |
| p25_fare | 119.08 | 37.71 | 147.01 | 39.54 |
| p50_fare | 149.53 | 49.33 | 181.89 | 49.93 |
| p75_fare | 208.09 | 79.69 | 239.68 | 69.50 |
| p85_fare | 256.43 | 113.94 | 285.68 | 89.74 |

*Passenger weighted mean

Table 3: Basic Regressions

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
|---------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| | Nonstop | Connect | Nonstop | Connect | Nonstop | Connect |
| bagfee | | | -0.0265*** (0.00562) | -0.0235*** (0.00332) | -0.0265* (0.0147) | -0.0235 (0.0180) |
| leg_ns2 | -0.0345*** (0.0115) | | -0.0348*** (0.0115) | | -0.0348*** (0.00578) | |
| leg_ns3 | 0.0609* (0.0369) | | 0.0608* (0.0369) | | 0.0608*** (0.00729) | |
| D(WN_ns) | -0.312*** (0.0140) | | -0.312*** (0.0140) | | -0.312*** (0.0107) | |
| D(otherlcc_ns) | -0.162*** (0.0121) | | -0.163*** (0.0122) | | -0.163*** (0.00855) | |
| leg_connect2 | | 0.0208*** (0.00546) | | 0.0207*** (0.00546) | | 0.0207*** (0.00498) |
| leg_connect3 | | -0.0273*** (0.00426) | | -0.0274*** (0.00426) | | -0.0274*** (0.00339) |
| leg_connect4m | | -0.00609 (0.00435) | | -0.00614 (0.00435) | | -0.00614** (0.00262) |
| D(WN_connect) | 0.00599 (0.0184) | -0.0590*** (0.00881) | 0.00577 (0.0184) | -0.0588*** (0.00882) | 0.00577 (0.00922) | -0.0588*** (0.00549) |
| D(otherlcc_connect) | 0.00230 (0.0202) | -0.101*** (0.00629) | 0.00255 (0.0202) | -0.101*** (0.00629) | 0.00255 (0.0109) | -0.101*** (0.00563) |
| D(WN_adj_ns) | -0.165*** (0.0164) | 0.0392 (0.0320) | -0.165*** (0.0164) | 0.0393 (0.0320) | -0.165*** (0.00905) | 0.0393*** (0.0141) |
| D(otherlcc_adj_ns) | -0.0663*** (0.0196) | 0.00228 (0.0216) | -0.0663*** (0.0196) | 0.00218 (0.0215) | -0.0663*** (0.00734) | 0.00218 (0.00719) |
| ltdist | 0.317*** (0.0102) | 0.267*** (0.00599) | 0.316*** (0.0102) | 0.267*** (0.00599) | 0.316*** (0.00743) | 0.267*** (0.00764) |
| D(connect) | -0.00452 (0.00649) | | -0.00466 (0.00648) | | -0.00466 (0.0110) | |
| pop | 0.000749 (0.00350) | -0.00960*** (0.00247) | 0.000769 (0.00350) | -0.00962*** (0.00247) | 0.000769 (0.00115) | -0.00962*** (0.00128) |
| income | 0.00441*** (0.00147) | -0.00142** (0.000576) | 0.00440*** (0.00148) | -0.00141** (0.000576) | 0.00440*** (0.00117) | -0.00141*** (0.000344) |
| tempdiff | -0.00344*** (0.000440) | -0.00312*** (0.000236) | -0.00343*** (0.000440) | -0.00312*** (0.000236) | -0.00343*** (0.000349) | -0.00312*** (0.000220) |
| d08q2 | 0.0131*** (0.00350) | 0.0307*** (0.00235) | 0.0144*** (0.00353) | 0.0321*** (0.00236) | 0.0144 (0.0149) | 0.0321** (0.0137) |
| d08q3 | 0.0446*** (0.00523) | 0.0428*** (0.00304) | 0.0550*** (0.00588) | 0.0531*** (0.00347) | 0.0550*** (0.0152) | 0.0531** (0.0211) |
| d08q4 | 0.0133*** (0.00515) | 0.0103*** (0.00289) | 0.0284*** (0.00641) | 0.0248*** (0.00373) | 0.0284 (0.0183) | 0.0248 (0.0227) |
| d09q1 | -0.0596*** (0.00515) | -0.0995*** (0.00303) | -0.0407*** (0.00697) | -0.0808*** (0.00422) | -0.0407** (0.0171) | -0.0808*** (0.0235) |
| d09q2 | -0.107*** (0.00519) | -0.100*** (0.00321) | -0.0879*** (0.00700) | -0.0814*** (0.00433) | -0.0879*** (0.0176) | -0.0814*** (0.0229) |
| d09q3 | -0.0820*** (0.00629) | -0.0785*** (0.00374) | -0.0615*** (0.00786) | -0.0596*** (0.00457) | -0.0615*** (0.0197) | -0.0596** (0.0235) |
| d09q4 | -0.0363*** (0.00524) | -0.0373*** (0.00346) | -0.0157** (0.00730) | -0.0184*** (0.00432) | -0.0157 (0.0179) | -0.0184 (0.0230) |
| Constant | 3.104*** (0.0684) | 3.640*** (0.0411) | 3.112*** (0.0686) | 3.646*** (0.0411) | 3.112*** (0.0703) | 3.646*** (0.0633) |
| Observations | 65,554 | 117,705 | 65,554 | 117,705 | 65,554 | 117,705 |
| Adjusted R^2 | 0.734 | 0.528 | 0.734 | 0.528 | 0.734 | 0.528 |

*** p<0.01, ** p<0.05, * p<0.1

Additional competition variables (see Table A1) and carrier fixed effects suppressed

Standard errors clustered by market (columns 1-4) and carrier-quarter (columns 5 and 6)

Table 4: Regressions with Interactions

| VARIABLES | (1) Nonstop | (2) Nonstop | (3) Nonstop | (4) Nonstop | (5) Connect | (6) Connect | (7) Connect | (8) Connect |
|-------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| bagfee | -0.0265*** (0.00568) | -0.0506*** (0.00718) | -0.0242*** (0.00585) | -0.0428*** (0.00823) | -0.0231*** (0.00334) | -0.0257*** (0.00343) | -0.0310*** (0.00343) | -0.0305*** (0.00350) |
| bagfee_tempdiff | -0.00115** (0.000501) | | | -0.00133*** (0.000473) | -0.000167 (0.000221) | | | -0.000196 (0.000220) |
| bagfee_income | | 0.00465*** (0.000986) | | 0.00358*** (0.00111) | | 0.00166*** (0.000516) | | -0.000127 (0.000555) |
| bagfee_dist | | | 2.22e-05** (9.73e-06) | 2.21e-05** (1.06e-05) | | | 3.54e-05*** (3.63e-06) | 3.58e-05*** (3.99e-06) |
| leg_ns2 | -0.0346*** (0.0115) | -0.0356*** (0.0114) | -0.0347*** (0.0114) | -0.0351*** (0.0113) | | | | |
| leg_ns3 | 0.0604 (0.0370) | 0.0615* (0.0365) | 0.0605* (0.0362) | 0.0606* (0.0360) | | | | |
| D(WN_ns) | -0.311*** (0.0140) | -0.311*** (0.0140) | -0.312*** (0.0139) | -0.311*** (0.0140) | | | | |
| D(otherlcc_ns) | -0.163*** (0.0122) | -0.163*** (0.0120) | -0.162*** (0.0121) | -0.162*** (0.0120) | | | | |
| leg_connect2 | | | | | 0.0207*** (0.00547) | 0.0208*** (0.00546) | 0.0227*** (0.00544) | 0.0226*** (0.00544) |
| leg_connect3 | | | | | -0.0274*** (0.00426) | -0.0275*** (0.00426) | -0.0286*** (0.00428) | -0.0285*** (0.00428) |
| leg_connect4m | | | | | -0.00613 (0.00435) | -0.00635 (0.00433) | -0.00849* (0.00434) | -0.00848* (0.00434) |
| D(WN_connect) | 0.00531 (0.0186) | 0.00403 (0.0186) | 0.00496 (0.0181) | 0.00310 (0.0184) | -0.0589*** (0.00881) | -0.0590*** (0.00882) | -0.0586*** (0.00876) | -0.0586*** (0.00875) |
| D(otherlcc_connect) | 0.00259 (0.0202) | 0.00412 (0.0199) | 0.00397 (0.0201) | 0.00523 (0.0199) | -0.101*** (0.00630) | -0.101*** (0.00629) | -0.1000*** (0.00628) | -0.0999*** (0.00629) |
| D(WN_adj_ns) | -0.165*** (0.0164) | -0.165*** (0.0162) | -0.163*** (0.0162) | -0.164*** (0.0161) | 0.0393 (0.0320) | 0.0382 (0.0319) | 0.0406 (0.0320) | 0.0407 (0.0320) |
| D(otherlcc_adj_ns) | -0.0662*** (0.0196) | -0.0658*** (0.0194) | -0.0655*** (0.0192) | -0.0648*** (0.0190) | 0.00224 (0.0215) | 0.00182 (0.0215) | 0.00144 (0.0215) | 0.00154 (0.0215) |
| ltdist | 0.316*** (0.0102) | 0.316*** (0.0101) | 0.306*** (0.0104) | 0.305*** (0.0106) | 0.267*** (0.00599) | 0.267*** (0.00599) | 0.245*** (0.00615) | 0.245*** (0.00619) |
| D(connect) | -0.00461 (0.00648) | -0.00499 (0.00642) | -0.00313 (0.00645) | -0.00333 (0.00643) | | | | |
| pop | 0.000831 (0.00350) | 0.000595 (0.00351) | 0.000564 (0.00340) | 0.000503 (0.00342) | -0.00962*** (0.00247) | -0.00971*** (0.00247) | -0.0103*** (0.00247) | -0.0103*** (0.00247) |
| income | 0.00446*** (0.00147) | 0.00225 (0.00148) | 0.00426*** (0.00147) | 0.00266* (0.00149) | -0.00141** (0.000576) | -0.00232*** (0.000611) | -0.00150*** (0.000575) | -0.00143** (0.000604) |
| tempdiff | -0.00285*** (0.000477) | -0.00346*** (0.000438) | -0.00339*** (0.000437) | -0.00274*** (0.000465) | -0.00303*** (0.000256) | -0.00312*** (0.000236) | -0.00313*** (0.000236) | -0.00302*** (0.000257) |
| d08q2 | 0.0148*** (0.00351) | 0.0148*** (0.00354) | 0.0144*** (0.00353) | 0.0152*** (0.00351) | 0.0321*** (0.00236) | 0.0320*** (0.00236) | 0.0315*** (0.00235) | 0.0315*** (0.00235) |
| d08q3 | 0.0554*** (0.00587) | 0.0555*** (0.00588) | 0.0553*** (0.00589) | 0.0561*** (0.00586) | 0.0531*** (0.00347) | 0.0529*** (0.00346) | 0.0523*** (0.00347) | 0.0523*** (0.00346) |
| d08q4 | 0.0290*** (0.00640) | 0.0287*** (0.00635) | 0.0284*** (0.00643) | 0.0293*** (0.00635) | 0.0248*** (0.00373) | 0.0245*** (0.00372) | 0.0236*** (0.00373) | 0.0236*** (0.00373) |
| d09q1 | -0.0397*** (0.00702) | -0.0393*** (0.00679) | -0.0403*** (0.00692) | -0.0381*** (0.00681) | -0.0807*** (0.00422) | -0.0810*** (0.00421) | -0.0817*** (0.00421) | -0.0817*** (0.00420) |
| d09q2 | -0.0873*** (0.00699) | -0.0870*** (0.00687) | -0.0876*** (0.00697) | -0.0862*** (0.00684) | -0.0814*** (0.00433) | -0.0816*** (0.00432) | -0.0822*** (0.00432) | -0.0823*** (0.00432) |
| d09q3 | -0.0612*** (0.00784) | -0.0609*** (0.00781) | -0.0611*** (0.00786) | -0.0603*** (0.00779) | -0.0596*** (0.00458) | -0.0597*** (0.00457) | -0.0601*** (0.00458) | -0.0601*** (0.00458) |
| d09q4 | -0.0152** (0.00728) | -0.0150** (0.00720) | -0.0153** (0.00728) | -0.0142** (0.00717) | -0.0184*** (0.00432) | -0.0185*** (0.00432) | -0.0189*** (0.00432) | -0.0189*** (0.00432) |
| Constant | 3.102*** (0.0685) | 3.212*** (0.0697) | 3.188*** (0.0723) | 3.253*** (0.0737) | 3.644*** (0.0411) | 3.681*** (0.0419) | 3.808*** (0.0430) | 3.805*** (0.0430) |
| Observations | 65,554 | 65,554 | 65,554 | 65,554 | 117,705 | 117,705 | 117,705 | 117,705 |
| Adjusted R ² | 0.734 | 0.735 | 0.735 | 0.736 | 0.528 | 0.528 | 0.530 | 0.530 |

*** p<0.01, ** p<0.05, * p<0.1

Additional competition variables and carrier fixed effects suppressed

Standard errors clustered by market in parentheses

Table 5: Regressions with Two Bag-Fee Variables

| VARIABLES | (1) | (2) |
|---------------------|---------------------------|---------------------------|
| | Nonstop | Connect |
| bagfee15 | -0.0248* (0.0146) | -0.0250 (0.0185) |
| bagfee20 | -0.0401** (0.0185) | -0.0156 (0.0206) |
| leg_ns2 | -0.0348*** (0.00578) | |
| leg_ns3 | 0.0609*** (0.00728) | |
| D(WN_ns) | -0.312*** (0.0107) | |
| D(otherlcc_ns) | -0.163*** (0.00856) | |
| leg_connect2 | | 0.0207*** (0.00497) |
| leg_connect3 | | -0.0274*** (0.00340) |
| leg_connect4m | | -0.00614** (0.00262) |
| D(WN_connect) | 0.00578 (0.00918) | -0.0589*** (0.00544) |
| D(otherlcc_connect) | 0.00253 (0.0109) | -0.101*** (0.00564) |
| D(WN_adj_ns) | -0.165*** (0.00904) | 0.0392*** (0.0141) |
| D(otherlcc_adj_ns) | -0.0663*** (0.00735) | 0.00222 (0.00719) |
| ltdist | 0.316*** (0.00743) | 0.267*** (0.00764) |
| D(connect) | -0.00477 (0.0110) | |
| pop | 0.000764 (0.00114) | -0.00962*** (0.00128) |
| income | 0.00440*** (0.00117) | -0.00141*** (0.000344) |
| tempdiff | -0.00343*** (0.000349) | -0.00312*** (0.000220) |
| d08q2 | 0.0144 (0.0147) | 0.0321** (0.0138) |
| d08q3 | 0.0545*** (0.0151) | 0.0538** (0.0212) |
| d08q4 | 0.0274 (0.0184) | 0.0258 (0.0229) |
| d09q1 | -0.0418** (0.0171) | -0.0795*** (0.0238) |
| d09q2 | -0.0891*** (0.0176) | -0.0802*** (0.0233) |
| d09q3 | -0.0595*** (0.0196) | -0.0606** (0.0239) |
| d09q4 | -0.00852 (0.0184) | -0.0226 (0.0231) |
| Constant | 3.113*** (0.0703) | 3.645*** (0.0635) |
| Observations | 65,554 | 117,705 |
| Adjusted R^2 | 0.734 | 0.528 |

*** p<0.01, ** p<0.05, * p<0.1

Additional competition variables and carrier fixed effects suppressed
Standard errors clustered by carrier-quarter in parentheses

Table 6: Percentile Regressions

| VARIABLES | (1) | (2) | (3) |
|---------------------|---------------------------|---------------------------|---------------------------|
| | Nonstop | Nonstop | Nonstop |
| bagfee_p25 | -0.0547*** (0.0197) | -0.0436** (0.0175) | -0.0506** (0.0203) |
| bagfee_p50 | | -0.0391** (0.0179) | |
| bagfee_p75 | | | -0.0354** (0.0173) |
| bagfee_p85 | -0.00104 (0.0164) | 0.0101 (0.0172) | 0.00314 (0.0159) |
| leg_ns2 | -0.0316*** (0.00676) | -0.0331*** (0.00644) | -0.0371*** (0.00677) |
| leg_ns3 | 0.0483*** (0.00911) | 0.0526*** (0.00916) | 0.0424*** (0.00897) |
| D(WN_ns) | -0.318*** (0.0118) | -0.309*** (0.0116) | -0.322*** (0.0124) |
| D(otherlcc_ns) | -0.175*** (0.00902) | -0.173*** (0.00890) | -0.179*** (0.00951) |
| D(WN_connect) | 0.00726 (0.00937) | 0.00920 (0.00956) | 0.00551 (0.00939) |
| D(otherlcc_connect) | 0.00231 (0.0107) | -0.00235 (0.00996) | 0.00514 (0.0114) |
| D(WN_adj_ns) | -0.173*** (0.00995) | -0.163*** (0.00978) | -0.175*** (0.0105) |
| D(otherlcc_adj_ns) | -0.0708*** (0.00826) | -0.0652*** (0.00767) | -0.0696*** (0.00871) |
| ltdist | 0.303*** (0.00783) | 0.290*** (0.00775) | 0.287*** (0.00786) |
| D(connect) | -0.00140 (0.0114) | 0.0137 (0.0114) | 0.00195 (0.0114) |
| pop | -0.000876 (0.00133) | -0.00232* (0.00119) | 0.000356 (0.00130) |
| income | 0.00399*** (0.00116) | 0.00347*** (0.00113) | 0.00443*** (0.00121) |
| tempdiff | -0.00342*** (0.000357) | -0.00339*** (0.000349) | -0.00366*** (0.000392) |
| d08q2 | 0.0173 (0.0155) | 0.0200 (0.0151) | 0.0106 (0.0152) |
| d08q3 | 0.0686*** (0.0153) | 0.0757*** (0.0150) | 0.0518*** (0.0150) |
| d08q4 | 0.0348* (0.0187) | 0.0405** (0.0178) | 0.0199 (0.0183) |
| d09q1 | -0.0479*** (0.0176) | -0.0517*** (0.0165) | -0.0512*** (0.0170) |
| d09q2 | -0.0867*** (0.0180) | -0.0926*** (0.0170) | -0.101*** (0.0173) |
| d09q3 | -0.0529** (0.0203) | -0.0561*** (0.0192) | -0.0692*** (0.0198) |
| d09q4 | -0.00936 (0.0186) | -0.0102 (0.0173) | -0.0224 (0.0184) |
| p25 | | -0.223*** (0.00544) | -0.536*** (0.0127) |
| p85 | 0.709*** (0.0182) | 0.486*** (0.0157) | 0.173*** (0.00782) |
| Constant | 2.828*** (0.0728) | 3.146*** (0.0694) | 3.470*** (0.0745) |
| Observations | 107,050 | 160,575 | 160,575 |
| Adjusted R^2 | 0.826 | 0.798 | 0.790 |

*** p<0.01, ** p<0.05, * p<0.1

Additional competition variables and carrier fixed effects suppressed
Standard errors clustered by carrier-quarter in parentheses

Table A1: Additional Competition Variables
in the Basic Model

| VARIABLES | (1) Nonstop | (2) Connect | (3) Nonstop | (4) Connect | (5) Nonstop | (6) Connect |
|-------------------------|------------------------|-------------------------|------------------------|-------------------------|-------------------------|-------------------------|
| nleg_connect | 0.00810 (0.00783) | | 0.00823 (0.00784) | | 0.00823** (0.00347) | |
| nleg_adj_ns | -0.0125 (0.00956) | 0.0175** (0.00885) | -0.0124 (0.00957) | 0.0175** (0.00885) | -0.0124** (0.00501) | 0.0175*** (0.00280) |
| nleg_adj_connect | | 0.0204*** (0.00708) | | 0.0204*** (0.00706) | | 0.0204*** (0.00400) |
| D(WN_adj_connect) | | -0.0308* (0.0181) | | -0.0317* (0.0179) | | -0.0317** (0.0130) |
| D(otherlcc_adj_connect) | | -0.0418** (0.0173) | | -0.0414** (0.0174) | | -0.0414*** (0.0157) |
| D(WN_pot) | -0.126*** (0.0179) | -0.0857*** (0.00903) | -0.126*** (0.0180) | -0.0858*** (0.00904) | -0.126*** (0.00896) | -0.0858*** (0.00604) |
| D(otherlcc_pot) | 0.0265 (0.0489) | -0.0691*** (0.0204) | 0.0267 (0.0488) | -0.0693*** (0.0203) | 0.0267** (0.0121) | -0.0693*** (0.00954) |
| D(WN_adj_pot) | -0.0554*** (0.0194) | 0.00568 (0.0128) | -0.0554*** (0.0194) | 0.00579 (0.0128) | -0.0554*** (0.00910) | 0.00579 (0.00725) |
| D(otherlcc_adj_pot) | -0.0542** (0.0211) | 0.00141 (0.0290) | -0.0542*** (0.0210) | 0.000894 (0.0290) | -0.0542*** (0.00832) | 0.000894 (0.00848) |

*** p<0.01, ** p<0.05, * p<0.1

Standard errors clustered by market (columns 1-4) and carrier-quarter (columns 5 and 6)

References

- ADAMS, W.J., YELLEN, J.L., 1976. Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics* 90, 475-498.
- ANGRIST, J., PISCHKE, J.-S., 2009. *Mostly Harmless Econometrics*. Princeton University Press, Princeton.
- BERTINI, M., OFEK, E., ARIELY, D., 2008. The impact of add-on features on consumer product evaluations. *Journal of Consumer Research* 36, 17-28.
- BORENSTEIN, S., 1989. Hubs and high fares: Dominance and market power in the U.S. airline industry. *RAND Journal of Economics* 20, 344-365.
- BRUECKNER, J.K., LEE, D.N., SINGER, E., 2012. Airline competition and domestic U.S. airfares: A comprehensive reappraisal. *Economics of Transportation* 2, 1-17.
- BRUECKNER, J.K., DYER, N., SPILLER, P.T., 1992. Fare determination in airline hub and spoke networks. *RAND Journal of Economics* 23, 309-333.
- BUREAU OF TRANSPORTATION STATISTICS, 2010. "Bag Fees by Airline." http://www.bts.gov/programs/airline_information/baggage_fees/.
- DAVIS S., MURPHY K., 2000. A competitive perspective on the Internet. *American Economic Review* 90, 184-7.
- EVANS, D.S., SALINGER, M.A., 2008. The role of cost in determining when firms offer bundles, *Journal of Industrial Economics* 41, 143-168.
- EVANS, D.S., SALINGER, M., 2005. Why do firms bundle and tie? Evidence from competitive markets and implications for tying law. *Yale Journal on Regulation* 22, 37-89.
- ELLISON, G., 2005. A model of add-on pricing. *Quarterly Journal of Economics* 120, 585-637.
- FRUCHTER, G.E., GERTSNER, E., DOBSON, P., 2011. Fee or free? How much to add on for an add-on? *Marketing Letters* 22, 65-88.
- GOOLSBEE, A., SYVERSON, C., 2008. How do incumbents respond to the threat of entry? Evidence from major airlines. *Quarterly Journal of Economics* 123, 1611-1633.
- GAYLE, P., WU, C.-Y., 2011. A re-examination of incumbents' response to the threat of entry: Evidence from the airline industry. Unpublished working paper, Kansas State

University.

- LEWBEL A., 1985. Bundling of substitutes or complements. *International Journal of Industrial Organization* 3, 101-7.
- MORRISON, S.A., 2001. Actual, adjacent, and potential competition: Estimating the full effect of Southwest Airlines. *Journal of Transport Economics and Policy* 32, 239-256.
- MORRISON, S.A., WINSTON, C., 1995. *Evolution of the Airline Industry*. Washington, D.C.: Brookings Institution.
- MORRISON, S.A., WINSTON, C., 1986. *The Economic Effects of Airline Deregulation*. Washington, D.C.: Brookings Institution.
- MUSSA, M., ROSEN, S., 1978. Monopoly and product quality, *Journal of Economic Theory* 18, 301-317.
- POSNER, R., 1976. *Antitrust Law: An Economic Perspective*. Chicago: University of Chicago Press.
- SCHMALENSEE R., 1982. Commodity bundling by a single-product monopolist. *Journal of Law and Economics* 25, 67-71.
- STIGLER, G., 1963. United States vs. Loew's Inc.: A note on block booking, *U.S. Supreme Court Review*, 152-175
- VERBOVEN, F., 1999. Product line rivalry and market segmentation—With an application to automobile optional engine pricing. *Journal of Industrial Economics* 47, 299-425.
- WHINSTON M., 1990. Tying, foreclosure, and exclusion. *American Economic Review* 80, 837-860.

Footnotes

*A portion of the empirical work in this paper is based on analysis contained in an expert report by Darin Lee, who was retained by Delta Air Lines in connection with a class action litigation filed in U.S. Federal Court. However, the opinions expressed are entirely ours. We thank an associate editor and a number of referees for helpful comments.

¹In addition to Adams and Yellen (1976) and Stigler (1963), see Schmalensee (1982) and Lewbel (1985) for other early papers on bundling. The literature presents models where consumers demand either zero or one unit of different products. Firms can offer no bundles, pure bundles or mixed bundles according to whether they offer their products separately, jointly, or in both fashions. The literature emphasizes the result that monopolies have incentives to bundle products to discriminate between consumers. Evans and Salinger (2005, 2008) emphasize cost-savings from bundling as another force behind the practice. The related tie-in literature argues that a firm can leverage its profit by tying one product to another. Posner (1976) argues that monopolies have no such incentives because they unprofitably lose the flexibility to adapt some prices under a tie-in. However, Tirole (1988) shows that monopolies should set the prices of complementary products to lower levels than would two independent firms. Davis and Murphy (2000) use this argument to argue that Microsoft followed a profitable path in setting Internet Explorer's price to zero. Whinston (1990) demonstrates that tie-ins could deter entry in a dynamic setting where they are irreversible.

²See, for example, "More consumers pack lighter, smarter to save when flying", *USA Today*, June 1, 2010.

³Fees for second checked bags were adopted prior to those for first checked bags. However, only a small share domestic passengers checks two pieces of luggage.

⁴Some individual fare reductions, however, could exceed the amount of the bag fee, an outcome that is shown to be theoretically possible.

⁵For a different type of evidence, see the experimental results on consumer valuation of add-ons provided by Bertini, Ofek and Ariely (2008).

⁶Some passengers could have $\delta < 0$, indicating a gain from not checking a bag (by avoiding, for example, the usual baggage-claim wait). The analysis would require some modification to handle this case.

⁷Because of a faster boarding process with fewer carry-on bags and a resulting increase in aircraft utilization, the per-passenger cost c could be an increasing function of the share of passengers checking bags. Taking such an effect into account, however, would complicate

the analysis.

⁸The second-order conditions for this maximization problem, which are not guaranteed to hold, are assumed to be satisfied.

⁹Note that continuity of the density is implicitly used in this argument, with $f(0) > 0$ implying that f is positive just above zero. Also, observe that the same conclusion would hold if the bottom of the support of f , denoted $\underline{\delta}$, were positive instead of zero. Raising the bag fee from zero up to $\underline{\delta}$ would have no effect since no one would pay it, but if $f(\underline{\delta}) > 0$, then the above argument shows the optimality of raising b above $\underline{\delta}$.

¹⁰Since the airline has more instruments at its disposal than in the pre-bag-fee situation, the profit it earns must be at least as high.

¹¹The proof is available on request.

¹²The proof, which is available on request, uses a comparative-statics approach, showing how the optimal p chosen conditional on b varies with b 's level. The results show that p declines with b when $b < k$. Therefore raising b from a value of zero reduces the fare, driving it below \tilde{p} , as long as b remains below k .

¹³Not being empirically oriented, neither of closest previous papers in the literature (Frutcher et al. (2011), Ellison (2005)) compares product prices before and after the unbundling of a product add-on. Therefore, Propositions 2 and 3 have no apparent precedent in the literature.

¹⁴See the United Parcel Service website at <http://www.ups.com/aircargo/using/rates/domestic.pdf>.

¹⁵For representative early papers, see Borenstein (1989), Brueckner, Dyer and Spiller (1992), and Morrison and Winston (1986, 1995). More recently, Morrison (2001) and Goolsbee and Syverson (2008) have studied the fare impacts of LCC competition.

¹⁶The legacy carriers are American, United, Delta, Continental, Northwest, US Airways, Alaska, and Midwest. The LCC carriers are Southwest, JetBlue, AirTran, Frontier, Virgin America, Spirit, Sun Country, and Allegiant.

¹⁷The list of sources (consisting of news articles and press releases) that give bag-fee imposition dates is available on request.

¹⁸The data set is limited to round-trip and one-way tickets for travel within the lower 48 US states; open-jaw itineraries (where a round-trip passenger does not return to the origin city) are dropped; interline itineraries (those with multiple “marketing” carriers) and itineraries with four or more coupons (flight segments) in either direction are dropped; itineraries with directional fares below \$25 are dropped; and itineraries with fares 5 times greater than the DOT’s Standard Industry Fare Level (“SIFL”) are excluded. In addition, the identity of carrier for the itinerary is its (single) marketing carrier, whose name appears on the ticket. While this convention automatically and appropriately identifies feeder carriers with their mainline partner, it also ignores codesharing between mainline carriers. In other words, mainline codeshare itineraries, where the mainline “operating” carrier for one or more route segments differs from the marketing carrier for that segment, are treated as equivalent to non-codeshare itineraries, where the operating and marketing carriers are the same for all segments.

¹⁹To avoid any bias resulting from the Delta/Northwest merger, Northwest fares are excluded from the sample. Independent competition from Northwest, however, is still measured in the pre-merger period. Since the time period covered by the BLS paper was entirely before the Delta/Northwest merger, this concern did not need to be addressed in that paper.

²⁰Markets with fewer than 10 passengers per day each way (“PPDEW”) are excluded from the nonstop sample, while markets with fewer than 5 PPDEW are excluded from the connect sample.

²¹Because the large number of competition variables in the model makes an instrumental-variable approach infeasible, the competition measures are treated as exogenous. It could be argued that any resulting simultaneity bias would reduce the absolute size of the competition effects, with the coefficients biased toward zero. The logic is that markets whose unobservable characteristics lead to high fares might be especially attractive for entry, leading to a positive correlation between the competition measures and the regression error term that biases the coefficients upward (toward zero). For several reasons, we think that any such bias is not likely to be substantial. First, Gayle and Wu (2011) show that simultaneity bias is negligible in a study that accounts for endogenous entry in airline markets. The second reason is the emergence of large LCC competition effects in the regressions, which would not emerge in the presence of substantial bias. Finally, even though the potential endogeneity of competition is based on plausible logic in the case of nonstop markets, the logic may not apply to connect markets. The reason is that a carrier’s presence in a connecting market is the result of separate decisions to provide service on the two spoke routes from its hub to the endpoint cities of the market. These service decisions in turn are based on overall network considerations, with traffic in potentially hundreds of markets involving one or the other of the spoke cities taken into account. With network considerations being paramount, favorable unobservables in a particular connecting market will have only a negligible influence on whether or not a carrier provides connecting service in that market. As result, there appears to be little basis for concern about correlation between the error term and the competition

measures in a fare regression for connect markets. Finally, note that any simultaneity bias in the competition coefficients is not likely to extend to the main coefficient of interest, that of the bag-fee dummy.

²²Only two airport-pair markets have more than three legacy competitors (Los Angeles-San Francisco and Los Angeles-Las Vegas), and these markets are dropped from the sample.

²³When the observation is for Southwest itself, $D(WN_ns)$ is set to zero to avoid counting Southwest as its own competitor. Similarly, when the observation is for another LCC, $D(otherlcc_ns)$ equals one only if another non-Southwest LCC different from the given LCC serves the market, being zero otherwise. Similar conventions apply to the remaining LCC competition dummy variables in both the nonstop and connect models.

²⁴In the underlying data, this variable takes a common value for all nonstop itineraries in a market, although its value may differ across connecting itineraries depending on the routing.

²⁵See Brueckner, Dyer and Spiller (1992) for more discussion of the density effect.

²⁶The presence of this variable means that, in the nonstop model, data is actually aggregated up to the market-carrier-quarter-connect level, with a carrier's connecting itineraries in a market separated from its nonstop itineraries. For simplicity, this point is glossed over in the text discussion.

²⁷It should be noted that the regression setup can be viewed as a type of difference-in-difference (DD) model. To understand this point, suppose that the data had a different structure, with only two carriers, one of which adopted a bag fee while the other did not, and only two time periods, one before and one after adoption of the fee. This structure would lead to the familiar DD regression, which would include a dummy variable for the bag-fee carrier, a dummy variable for the post-fee period, and the interaction of the two dummies. The present model differs by having more bag-fee carriers and more time periods, but it effectively retains the DD structure.

²⁸The DB1B indicates whether a coach fare has restrictions or is unrestricted (being fully refundable), but this distinction is not applied consistently by carriers in their reporting to the DOT.

²⁹When dummy coefficients in a regression are large, as is the $D(WN_ns)$ coefficient, the numerical value overstates the percentage impact on the dependent variable by a non-trivial factor. However, since the bag-fee coefficients that are of main interest have a smaller magnitude, making this overstatement less severe, the issue is ignored in the discussion.

³⁰The effects of distance and *tempdiff* are again positive and negative, respectively, and both *income* and *pop* now have significant effects, both of which are negative. Apparently, the airline cost reduction from higher traffic densities in larger, higher-income connect markets dominates the upward fare pressure from higher demand (seen in the nonstop model), leading to fare reductions. Given that connect markets are thinner than nonstop markets, allowing greater scope for a density effect, these findings make sense.

³¹Under a different approach, individual carriers would constitute groups for the purposes of clustering. In a purely cross-sectional model focusing on the effect of a carrier-specific variable, this grouping strategy would be appropriate, with the standard errors clustered by carrier to control for within-carrier error correlation. However, having an 8-quarter panel rather than a cross section, we can use carrier fixed effects, netting out one type of correlation that would be addressed by clustering: within-carrier random influences that are perfectly (or almost perfectly) correlated across quarters. Carrier-level clustering still serves to control for the effects of within-carrier shocks that are less correlated across quarters, as well as within-quarter shocks that are correlated across the markets served by the carrier. But because the *bagfee* variable varies by quarter, jumping from zero to a positive value once a bag fee is imposed (and thus changing the fare environment for the carrier), it appears that carrier-level clustering, which does not formally recognize this breakpoint, may be too stringent. As a result, we chose to cluster at the carrier-quarter level in all but one set of regressions, an approach that captures within-carrier, cross-market error correlation in each individual quarter. While this scheme does not capture within-carrier, cross-quarter correlation, we believe that this correlation is largely eliminated by the carrier fixed effect. Note also that the presence of quarter dummies further reduces concerns about such correlation, given that cross-quarter carrier shocks are mostly felt industry wide (fuel-price increases, for example), with these shocks being captured by the quarter dummies. For a recent exposition of the issues involved in clustering, see Angrist and Pischke (2009, Ch. 8).

³²While the coefficients of the quarterly dummies naturally change somewhat in response to the temporal pattern of the bag-fee variable, the coefficients of the other covariates are largely unaffected by its presence.

³³The following carriers raised their bag fees to \$20 during the sample period: Spirit (February 20, 2008), Delta (August 4, 2009), American (August 14, 2009), Continental (August 19, 2009), Virgin America (September 10, 2009), Frontier (October 1, 2009), and US Airways (October 7, 2009).

³⁴The number of observations is less than twice as large as in the earlier regressions because we remove observations where fare variety is insufficient to allow the 25th, 50th, 75th and 85th percentile fares to all be different (the number of observations increases by 63 percent).

³⁵When separate regressions are run, one for each of the four percentile levels shown in Table

6, the *bagfee* coefficient is only significant in the 25th percentile regression. Stacking the fare data yields greater precision in the estimates. Another advantage of stacking is that it allows significance tests on the differences between the bag-fee coefficients.